



1645

1600

RECEIVED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/765,061B

DATE: 02/04/2002  
TIME: 13:02:31

Input Set : A:\Copy of SQ-09765061.txt  
Output Set: N:\CRF3\02042002\I765061B.raw

5 FEB 13 2002  
TECH CENTER 1600/2900

2 <110> APPLICANT: Board of Regents of the University of Texas System  
4 <120> TITLE OF INVENTION: Mutations in a Novel Photoreceptor-pineal gene 17P cause  
5 leber congenital amaurosis (LCA4)  
7 <130> FILE REFERENCE: 96606/16UTL  
W--> 8 <140> CURRENT APPLICATION NUMBER: 09/765,061B  
9 <141> CURRENT FILING DATE: 2001-01-17  
11 <160> NUMBER OF SEQ ID NOS: 78  
12 <170> SOFTWARE: PatentIn version 3.1  
14 <210> SEQ ID NO: 1  
15 <211> LENGTH: 6689  
16 <212> TYPE: DNA  
17 <213> ORGANISM: Homo sapiens  
20 <220> FEATURE:  
21 <221> NAME/KEY: gene  
22 <222> LOCATION: (1)..(6689)  
23 <223> OTHER INFORMATION: the AIP1 gene produces aryl-hydrocarbon receptor  
24 interacting protein-like 1  
26 <220> FEATURE:  
27 <221> NAME/KEY: misc\_feature  
28 <222> LOCATION: (1897)..(1906)  
29 <223> OTHER INFORMATION: n represents any of the four nucleotides A T G or C  
31 <220> FEATURE:  
32 <221> NAME/KEY: misc\_feature  
33 <222> LOCATION: (3946)..(3946)  
34 <223> OTHER INFORMATION: n represents any of the four nucleotides A T G or C  
36 <400> SEQUENCE: 1  
37 ggctcccaa agtgcgtggat tacaggcggt agtcaccgcg cctggtcccc tgtttttttt 60  
39 aagaaagctc agcgacacctt ttcccttctt ggggtggaac aaaaagccaa atctagcaca 120  
41 accctggca ggggcccaga atcaactggaa gcaaagggtgg atgggatagg aggcgaggct 180  
43 gcctgtggac cacaggcccgg ccccgagtgg ctctgatgag aagccggggc gccttaggtca 240  
45 ccccccac cgtctgcct tccccccact cctcctggct gggtaaatcc cagagtctca 300  
47 gcccctaag tgtttttttt ggagggtgaga ttatctccgc ctgtgctgga caccccttt 360  
49 tctctgcag ccatggatgc cgctctgctc ctgaacgtgg aaggggtcaa gaaaaccatt 420  
51 ctgcacgggg gcacggcga gctcccaaacc ttcatcaccg gatcccgagt gagttggggcc 480  
53 cttccggagc agacagggtc cccacagca gctttcaaca ttccagggtgt gcccccaaggc 540  
55 actgtaaaca gctttcagct gtgcggaaaa aacagccagg cagccccagc gctgggcctc 600  
57 cggggagctc ccagcggtta cccattcagg gggcattttt ggtactttgc agattcaact 660  
59 ttagcatggg ctgagggaa gggctttgg gaattttctg gggccctaaa tggttagtga 720  
61 gaagaaaggg agtccgagga gtcttggat ttgtcccaa atgtctgtta ggcttccctg 780  
63 gactgaaggg tgcgtctgt gctacagaat tcgggctttg gccaggcggag gcccgtcccg 840  
65 cctgtaatcc cagcactttg ggaggccaag atgggcagat catgaggtca agagttcggag 900  
67 accagcctga ccaacatgtg aaaccccatc tctactgaaa atacaaaaat tagccagatg 960  
69 tgctgtggcg cctgtaatcc cagttcagat actcaggaga cttgaggcag gagaatcact 1020

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/765,061B

DATE: 02/04/2002  
TIME: 13:02:31

Input Set : A:\Copy of .SQ-09765061.txt  
Output Set: N:\CRF3\02042002\I765061B.raw

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/765,061B

DATE: 02/04/2002

TIME: 13:02:31

Input Set : A:\Copy of SQ-09765061.txt  
 Output Set: N:\CRF3\02042002\I765061B.raw

169	tgcatggAAC	tccccacATC	ctcccaacCCA	ggagccccAG	agtgattGGC	agcgcgtGTT	4020
171	tgtggattGG	tgagagaggG	ttagggccAG	ggtaaaggTC	aggtcaggAC	tcaagcttatG	4080
173	gccaagactG	aggctcagCC	tgagagctAT	gtgggtGAAT	aaaataaaAT	aagaactgtG	4140
175	tcaaccaaggG	gcccottACA	gcttgctGT	cacagttGT	tggctgtGc	actgcacaAG	4200
177	gtgcaccGGC	atctcctCCA	aggtgctAT	tatacacATT	gtatattGgt	attccataAA	4260
179	tgagaagttT	ccagcagatG	gcaatagtGT	attttctAA	caaaaacgAG	attcgtgaca	4320
181	attttctgAA	tattagaagtG	gaagtgtCTT	gatgaacGGG	cacctttCC	tagttgcac	4380
183	aaagacattG	attagggCA	gggtttcCG	cgttgttGCT	tctttccCTT	gtctgtatGC	4440
185	acttgaccAG	caagcatgAC	ttcaggggAGA	tgtgccacAG	ggtcctgtTT	ttcgggtCTC	4500
187	tgatgggtG	caggcccCTG	gggtccctGC	ctcaactgACC	tgcagctCTG	ggccaggtT	4560
189	gatgccccGA	gtgattacCA	gaggggagAC	tggaacctGA	gcaatcatGA	gaagatgaAG	4620
191	gcgggtGCCG	tcctccacCG	agagggaaAT	cggcttCA	agctgggCCG	ctacgaggAG	4680
193	gcctcttCCA	agtaccAGGA	gcccattcATC	tgcctaAGGA	acctgcAGAC	caaggtcAGA	4740
195	ggccgctGGC	caggggtGGG	aagtggcGT	gactctGGG	ggectgcccA	gtgccggCCA	4800
197	gggtggggCG	ggggttggGC	agctgcctGA	ggtcatggCT	gaccttctCC	ctgggcagGA	4860
199	gaagccatGG	gagggtgcAGT	ggctgaAGCT	ggagaagatG	atcaataACT	tgatcctCAA	4920
201	ctactgcccAG	tgcctgctGA	agaaggAGGA	gtactatGAG	tgctggAGC	acaccagtGA	4980
203	tattctccGG	caccacccAG	gtgcgcGGG	ctgcaggGGC	ggacagtGAG	ggggcgcccA	5040
205	gcccagggCC	acggagacAC	ctgcccataGC	cttcttggAC	ttttcttCC	accccaccAG	5100
207	ggcaccaAAAC	tttgtctCCA	cccagccGGG	tttccccGAG	tgtgtaaCTG	aattgtggGT	5160
209	gatggatGGG	cagtgttGG	cgcggggGGG	ccttattttA	aatgtgttGTT	tgaacactTA	5220
211	cccaggaAGG	tgcctaAGCT	tgtgatttCA	gccaacGGT	aaacaggcGT	ttaaaaAGAG	5280
213	gggcaatCAA	tataggAAA	aatattatGA	tgtcggtACT	agtactgtG	ttgcgaggAT	5340
215	atggcaccGC	agtactagAT	tgacttaATG	ctcaatcGT	gctcacAGTA	aaaacatCCA	5400
217	gcccctGGCT	catgcataCG	gcacacgtCG	tctgcgttTA	ttatctcATT	taatcctCAT	5460
219	aatcctcATA	atcaccatAT	gagggaggGT	cagggaaAGG	ggcctaaggG	ttatctaATT	5520
221	tagtgtcGT	ctataagAAA	aataaaACAA	agttatGAAT	ataaaaattAC	tcacagggCC	5580
223	ttaaaaAGGA	gaggaggAGG	tactgttATT	atgatcatCA	tctccatCTT	acagttGAGG	5640
225	aaaccgagggA	atggggataA	cagagaggTT	aaggatcatG	gcggggctGA	gggtcttGGA	5700
227	ggctgggtGA	tcccagctGG	gctggggcTG	cctctgaggC	tgggaaggGA	gctgtagCTG	5760
229	gatgctccCT	gctccccaca	ggcatactGA	aggctactA	cgtgcgtGCC	cgggctcacG	5820
231	cagagggtGT	gaatgaggCC	gaggccaAGG	cggacctCCA	gaaagtGCTG	gagctggAGC	5880
233	cgtccatGCA	gaaggcggtG	cgcaggGGAGC	ttgaggctGC	tggagaACCG	catggcgGAG	5940
235	aacaggagGA	ggagcggcTG	cgctgcccGA	acatgtcAG	ccagggTGCC	acgcagcCTC	6000
237	ccgcagagCC	acccacAGAG	ccacccgcAC	agtcatccAC	agagccacCT	gcagagccAC	6060
239	ccacagcaccA	atctgcAGAG	ctgtccgcAG	ggccccctGC	agagccAGCC	acagagccAC	6120
241	ccccgtcccc	agggcactCG	ctgcagcact	gagccccCTG	aggcccACAG	ccacccAGGC	6180
243	agggagcaAG	tggcctggTC	acttctggTT	cgattgacCA	ggatcgtGgt	gtcacttttt	6240
245	aaaatttAA	attaattttT	gaaatcaaAG	tcagacacAC	ccatggtaAA	aaaaaaaaAA	6300
247	aaaacaatCC	caagggtaca	gaagagCTA	tgaataaaAG	tagtttCTC	ctctacccCT	6360
249	ctcatccCT	ccgtccatG	gttttaATTG	accctgttt	taattcttCT	ggttagttTC	6420
251	tctatttCCA	agtaatctGT	ttaaatcAGT	ttctagattT	tacccatGT	caatgacaaa	6480
253	tgaggatttG	atgcctgtAT	ccttctcAT	gcctgatACC	cctccctgtC	tcccatttt	6540
255	ggatagttAC	atttggggGT	catctcggtG	attttgtAA	cttacgcAG	gacacttagA	6600
257	gctctctaga	atcccactGA	ctttagtGGG	gtcttgatGT	agggtgggCA	agccccgaca	6660
259	ctggagcttA	gcctgagaggG	ggttcttG				6689
263	<210> SEQ ID NO: 2						
264	<211> LENGTH: 1119						
265	<212> TYPE: DNA						

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/765,061B

DATE: 02/04/2002  
TIME: 13:02:31

Input Set : A:\Copy of SQ-09765061.txt  
Output Set: N:\CRF3\02042002\I765061B.raw

266 <213> ORGANISM: Papio anubis  
 268 <220> FEATURE:  
 269 <221> NAME/KEY: gene  
 270 <222> LOCATION: (1)..(1119)  
 271 <223> OTHER INFORMATION: the AIPL1 gene produces aryl-hydrocarbon receptor  
 272 interacting protein-like 1  
 274 <400> SEQUENCE: 2  
 276 atggatgccg ctctgctctt gaacgtggaa ggggtcaaga aaaccattct gcacggaggc 60  
 278 acgggcgagc tcccaaactt catcaccgga tcccgagtga tctttcattt ccgcaccatg 120  
 280 aaatgtatg aggagcgcac ggtcatcgac gacagccgac aggtggacca gcccatgcac 180  
 282 atcatcatcg ggaacatgtt caagctcgag gtctgggaga tcctgctcac ctccatgagg 240  
 284 gtgcacgagg tggccgagtt ctgggtgcac accatccaca cgggggtcta ccccatcctg 300  
 286 tcccgagcc tgcggcagat gcccggggc aaggacccca cggagtggca cgtgcacaca 360  
 288 tgcgggctgg ccaacatgtt cgcctaccac acactggct acgaggacct ggacgagctg 420  
 290 cagaaggagc ctcagctctt gatcttttgatc gagactgc tgcaggttga cgccccgagt 480  
 292 gattaccaga gggagacctg gaacctgagc aatcatgaga agatgaaggt ggtgcccgtc 540  
 294 ctccacggag agggaaatcg gctttcaag ctggggcgtc acgaggagggc cttttccaag 600  
 296 taccaggagg ccatcatctg cctaaggaac ctgcagacca aggagaagcc atgggaggtg 660  
 298 cagtggctga agctggagaa gatgatcaac accctgaccc tcaactactg ccagtgcctg 720  
 300 ctgaagaagg aggagattt cgggggtctg gagcacacca gtgacattt ccgcaccac 780  
 302 ccaggcatcg tgaaggccctt ctatgtgcgt gcccgggctc acgcagaggt gtgaatgag 840  
 304 gcccggggc aaggcgaccc ccaaaaatgtt ctggagctgg agccatccat gcaaggcg 900  
 306 gtgcgcgggg agctggaggt gctggagaac cgcattgcag agaaggcaga ggaggagcgg 960  
 308 ctgcgtgcctt ggaacatgtt gagccggggc gccacgcac cttccacacca gcccacggca 1020  
 310 gagcccccaca cagcaccacc tgcggagctg tccacaggc cacctgcaga gcccacccgca 1080  
 312 gagctcccccc tgcggggcagg gcactcactg cagcaactga 1119  
 316 <210> SEQ ID NO: 3  
 317 <211> LENGTH: 1155  
 318 <212> TYPE: DNA  
 319 <213> ORGANISM: Pan troglodytes  
 321 <220> FEATURE:  
 322 <221> NAME/KEY: gene  
 323 <222> LOCATION: (1)..(1155)  
 324 <223> OTHER INFORMATION: the AIPL1 gene produces aryl-hydrocarbon receptor  
 325 interacting protein-like 1  
 327 <400> SEQUENCE: 3  
 328 atggatgccg ctctgctctt gaacgtggaa ggggtcaaga aaaccattct gcacgggggc 60  
 330 acgggcgagc tcccaaactt catcaccgga tcccgagtga tctttcattt ccgcaccatg 120  
 332 aaatgtatg aggagcggac agtcattgcac gacagccgac aggtggggca gcccatgcac 180  
 334 atcatcatcg gaaacatgtt caagctcgag gtctgggaga tcctgctcac ctccatgcgg 240  
 336 gtgcacgagg tggccgagtt ctgggtgcac accatccaca cgggggtcta ccccatcctg 300  
 338 tcccgagcc tgaggcagat gcccggggc aaggacccca cagagtggca cgtgcacaca 360  
 340 tgcgggctgg ccaacatgtt cgcctaccac acgctggct acgaggacct ggacgagctg 420  
 342 cagaaggagc ctcagctctt gtttttttgc atcgagactgc tgcaggttga tgccccgagt 480  
 344 gattaccaga gggagacctg gaacctgagc aatcatgaga agatgaaggc ggtgcccgtc 540  
 346 ctccacgggg agggaaatcg gctttcaag ctggggcgtc acgaggagggc cttttccaag 600  
 348 taccaggagg ccatcatctg cctaaggaac ctgcagacca aggagaagcc gtggggaggtg 660  
 350 cagtggctga agctggagaa gatgatcaat actctgtatcc tcaactactg ccagtgcctg 720  
 352 ctgaagaagg aggagacta tgagggtctg gagcacacca ggcacattt ccggcaccac 780

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/765,061B

DATE: 02/04/2002  
TIME: 13:02:31

Input Set : A:\Copy of SQ-09765061.txt  
Output Set: N:\CRF3\02042002\I765061B.raw

354 ccaggcatcg tgaaggccta ctacgtgcgt gccccggc acgcagaggt gtggaatgag 840  
 356 gccgaggcca aggcagaccc ctggaaagtg ctggagctgg agccgtccat gcagaaggcg 900  
 358 gtgcgcaggg agctgaggct gctggagaac cgcatgggg agaagcagga ggaggagcgg 960  
 360 ctgcgcgtgcc ggaacatgct gagccagggt gccacgcgc ctccggcaga gccaccacaca 1020  
 362 gagccaccgg cacagtcatc cacagagcca cctgcagagc caccggcacc accatctgca 1080  
 364 gagctgtccg cagggccacc tgcagagaca gccacagagc caccggcgtc cccagggcac 1140  
 366 tcgcgtgcaggc actga 1155 -  
 370 <210> SEQ ID NO: 4  
 371 <211> LENGTH: 1060  
 372 <212> TYPE: DNA  
 373 <213> ORGANISM: Bos taurus  
 375 <220> FEATURE:  
 376 <221> NAME/KEY: gene  
 377 <222> LOCATION: (1)..(1060)  
 378 <223> OTHER INFORMATION: the AIPL1 gene produces aryl-hydrocarbon receptor  
 379 interacting protein-like 1  
 381 <400> SEQUENCE: 4  
 382 atggatgcca ctctgctcct gaatgtggaa gggatcaaga aaaccattct gcatggggc 60  
 384 acaggggacc tccccaaactt cattactgga gcccggatgt ccttcattt ccgaaccatg 120  
 386 aaatgtgatg aggagcggac ggtgatagac gacagcaagc aggtggggca tcccatgcac 180  
 388 atcatcattt ggaacatgtt caagctggag gtctggaga tcttgctgac gtcgcacac 240  
 390 gtcagcgagg tggccgagtt ttgggtgcac accatccaca caggggtcta cccatcctg 300  
 392 tcccgaggcc tgcggcagat ggcggagggt aaggacccca cagagtggca cgtgcacac 360  
 394 tggcgttgg ccaacatgtt cgcttaccac acgctgggt acgaggaccc ggacgagctg 420  
 396 cagaaggagc ctcagccact gatcttcata atcgagtgc tgcaggtcga ggcggcggc 480  
 398 cagtaccaga gggagacctg gaacctgaat aaccaggaga agatgcaggc ggtggccatc 540  
 400 ctccatggag aaggaaaccg gctcttcaag ctggggccct acgaggaccc ctccaacaag 600  
 402 taccaggaag ccatcgctg cttggaggaac ctgcagacca aggagaaacc ctggggagg 660  
 404 cagtggctga agctggagaa gatgatcaac accctgatcc tgaactactg tcagtgtctg 720  
 406 ctgaagaagg aggagacta cgagggtctg gaacacacta gtgacatctt ccggcatcac 780  
 408 ccaggcatcg tgaaggccta ctatgtgagg gccccggc acgcccgggt gtggaatgag 840  
 410 gcgaaagcca aggcggatct ggagaaagtg ctggagctgg agccgtccat gcgaaaggcg 900  
 412 gtgcagaggg agctgaggct gctggagaac cggctggagg agaaacgcga ggaggagcga 960  
 414 ctgcgcgtgcc ggaacatgct gggctagtgc gcaggcgcca agcctcctgc ctccggcc 1020  
 416 cgcycctcca ccccccacaa aaaaaaaaaa aaaaattttt 1060  
 419 <210> SEQ ID NO: 5  
 420 <211> LENGTH: 925  
 421 <212> TYPE: DNA  
 422 <213> ORGANISM: Canis familiaris  
 424 <220> FEATURE:  
 425 <221> NAME/KEY: gene  
 426 <222> LOCATION: (1)..(925)  
 427 <223> OTHER INFORMATION: the AIPL1 gene produces aryl-hydrocarbon receptor  
 428 interacting protein-like 1  
 430 <400> SEQUENCE: 5  
 431 tgtacggggg caccggcgag ctcccaaactt tcctcacggg gtccgggtc atcttcact 60  
 433 tccgcacaaac gaaatgcgcac gaggcgccga cagtgtgcga cgacagcaag cgtgtggcc 120  
 435 atcccatgca catcatcatac ggaaacatgt tcaagctggaa ggtctggag gtgtgtctg 180  
 437 catccatgcg cgtggcgag gtggccgagt tctgggtgcga ctctattcac acaggagtct 240

Use of n and/or Xaa has been detected in the Sequence Listing.  
 → Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/765,061B

DATE: 02/04/2002

TIME: 13:02:32

Input Set : A:\Copy of SQ-09765061.txt  
Output Set: N:\CRF3\02042002\I765061B.raw

L:8 M:283 W: Missing Blank Line separator, <140> field identifier

L:99 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:1717 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72